

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/811,088

DATE: 03/30/2001

TIME: 15:03:21

Input Set : A:\07334-324001.TXT

Output Set: N:\CRF3\03302001\I811088.raw

OIPE

RECEIVED

APR 11 2001

TECH CENTER 1600/2900

#2

4 <110> APPLICANT: Holtzman, Douglas A.  
 5 Gearing, David P.  
 6 Pan, Yang  
 8 <120> TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
 9 PROGNASTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER  
 10 USES  
 12 <130> FILE REFERENCE: 07334-324001  
 C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/811,088  
 C--> 14 <141> CURRENT FILING DATE: 2001-03-16

ENTERED

14 <150> PRIOR APPLICATION NUMBER: US 09/712,726  
 15 <151> PRIOR FILING DATE: 2000-11-14  
 17 <150> PRIOR APPLICATION NUMBER: US 08/820,364  
 18 <151> PRIOR FILING DATE: 1997-03-12  
 20 <150> PRIOR APPLICATION NUMBER: US 09/757,421  
 21 <151> PRIOR FILING DATE: 2001-01-10  
 23 <150> PRIOR APPLICATION NUMBER: US 08/843,652  
 24 <151> PRIOR FILING DATE: 1997-04-16  
 26 <150> PRIOR APPLICATION NUMBER: US 08/843,651  
 27 <151> PRIOR FILING DATE: 1997-04-16  
 29 <150> PRIOR APPLICATION NUMBER: US 09/354,809  
 30 <151> PRIOR FILING DATE: 1999-07-16  
 32 <150> PRIOR APPLICATION NUMBER: US 08/938,365  
 33 <151> PRIOR FILING DATE: 1997-09-26  
 35 <160> NUMBER OF SEQ ID NOS: 24  
 37 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 39 <210> SEQ ID NO: 1  
 40 <211> LENGTH: 755  
 41 <212> TYPE: DNA  
 42 <213> ORGANISM: Homo sapiens  
 44 <220> FEATURE:  
 45 <221> NAME/KEY: CDS  
 46 <222> LOCATION: (1)...(279)  
 48 <400> SEQUENCE: 1

49	atg gct cgc cta cag act gca ctc ctg gtt gtc ctc gtc ctc ctt gct	48
50	Met Ala Arg Leu Gln Thr Ala Leu Leu Val Val Leu Val Leu Leu Ala	
51	1 5 10 15	
53	gtg gcg ctt caa gca act gag gca ggc ccc tac ggc gcc aac atg gaa	96
54	Val Ala Leu Gln Ala Thr Glu Ala Gly Pro Tyr Gly Ala Asn Met Glu	
55	20 25 30	
57	gac agc gtc tgc tgc cgt gat tac gtc cgt tac cgt ctg ccc ctg cgc	144
58	Asp Ser Val Cys Cys Arg Asp Tyr Val Arg Tyr Arg Leu Pro Leu Arg	
59	35 40 45	
61	gtg gtg aaa cac ttc tac tgg acc tca gac tcc tgc ccg agg cct ggc	192
62	Val Val Lys His Phe Tyr Trp Thr Ser Asp Ser Cys Pro Arg Pro Gly	
63	50 55 60	
65	gtg gtg ttg cta acc ttc agg gat aag gag atc tgt gcc gat ccc aga	240
66	Val Val Leu Leu Thr Phe Arg Asp Lys Glu Ile Cys Ala Asp Pro Arg	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/811,088

DATE: 03/30/2001

TIME: 15:03:24

Input Set : A:\07334-324001.TXT

Output Set: N:\CRF3\03302001\I811088.raw

```

67 65          70          75          80
69 gtg ccc tgg gtg aag atg att ctc aat aag ctg agc caa tgaagagcct      289
70 Val Pro Trp Val Lys Met Ile Leu Asn Lys Leu Ser Gln
71          85          90
73 actctgatga ccgtggcctt ggctcctcca ggaaggctca ggagccctac ctccctgccca      349
74 ttatagctgc tccccgccag aagcctgtgc caactctctg cattccctga tctccatccc      409
75 tgtggctgtc acccttggtc acctccgtgc tgtcactgcc atctccccc tgacccctct      469
76 aaccctaccc ctgcctccct ccctgcagtc agaggggtcct gttcccatca gcgattcccc      529
77 tgcttaaacc cttccatgac tccccactgc cctaagctga ggtcagttct ccaagcctgg      589
78 catgtggccc tctggatctg ggttccattt ctgtctccag cctgcccact tcccttcatg      649
79 aatgttgggt tctagctccc tgtttccaa acccatacta cacatccac ttctgggtct      709
80 ttgcctggga tgttctgac actcagaaag tcccgctgac gcggcc      755
82 <210> SEQ ID NO: 2
83 <211> LENGTH: 93
84 <212> TYPE: PRT
85 <213> ORGANISM: Homo sapiens
87 <400> SEQUENCE: 2
88 Met Ala Arg Leu Gln Thr Ala Leu Leu Val Val Leu Val Leu Leu Ala
89 1          5          10          15
90 Val Ala Leu Gln Ala Thr Glu Ala Gly Pro Tyr Gly Ala Asn Met Glu
91          20          25          30
92 Asp Ser Val Cys Cys Arg Asp Tyr Val Arg Tyr Arg Leu Pro Leu Arg
93          35          40          45
94 Val Val Lys His Phe Tyr Trp Thr Ser Asp Ser Cys Pro Arg Pro Gly
95          50          55          60
96 Val Val Leu Leu Thr Phe Arg Asp Lys Glu Ile Cys Ala Asp Pro Arg
97 65          70          75          80
98 Val Pro Trp Val Lys Met Ile Leu Asn Lys Leu Ser Gln
99          85          90
101 <210> SEQ ID NO: 3
102 <211> LENGTH: 4051
103 <212> TYPE: DNA
104 <213> ORGANISM: Homo sapiens
106 <220> FEATURE:
107 <221> NAME/KEY: CDS
108 <222> LOCATION: (128)...(1447)
110 <221> NAME/KEY: misc_feature
111 <222> LOCATION: (1)...(4051)
112 <223> OTHER INFORMATION: n = A,T,C or G
114 <400> SEQUENCE: 3
115 gtcgacccac gcgtccggcc ggagaacccg caatctttgc gccacaaaa tacaccgacg      60
116 atgcccgatc tactttaagg gctgaaaccc acgggcctga gagactataa gagcggtccc      120
117 taccgcc atg gaa caa cgg gga cag aac gcc ccg gcc gct tcg ggg gcc      169
118 Met Glu Gln Arg Gly Gln Asn Ala Pro Ala Ala Ser Gly Ala
119 1          5          10
121 cgg aaa agg cac ggc cca gga ccc agg gag gcg cgg gga gcc agg cct      217
122 Arg Lys Arg His Gly Pro Gly Pro Arg Glu Ala Arg Gly Ala Arg Pro
123 15          20          25          30
125 ggg ctc cgg gtc ccc aag acc ctt gtg ctc gtt gtc gcc gcg gtc ctg      265

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/811,088

DATE: 03/30/2001

TIME: 15:03:24

Input Set : A:\07334-324001.TXT

Output Set: N:\CRF3\03302001\I811088.raw

126	Gly	Leu	Arg	Val	Pro	Lys	Thr	Leu	Val	Leu	Val	Val	Ala	Ala	Val	Leu	
127					35				40						45		
129	ctg	ttg	gtc	tca	gct	gag	tct	gct	ctg	atc	acc	caa	caa	gac	cta	gct	313
130	Leu	Leu	Val	Ser	Ala	Glu	Ser	Ala	Leu	Ile	Thr	Gln	Gln	Asp	Leu	Ala	
131				50					55					60			
133	ccc	cag	cag	aga	gcg	gcc	cca	caa	caa	aag	agg	tcc	agc	ccc	tca	gag	361
134	Pro	Gln	Gln	Arg	Ala	Ala	Pro	Gln	Gln	Lys	Arg	Ser	Ser	Pro	Ser	Glu	
135				65					70					75			
137	gga	ttg	tgt	cca	cct	gga	cac	cat	atc	tca	gaa	gac	ggt	aga	gat	tgc	409
138	Gly	Leu	Cys	Pro	Pro	Gly	His	His	Ile	Ser	Glu	Asp	Gly	Arg	Asp	Cys	
139		80					85					90					
141	atc	tcc	tgc	aaa	tat	gga	cag	gac	tat	agc	act	cac	tgg	aat	gac	ctc	457
142	Ile	Ser	Cys	Lys	Tyr	Gly	Gln	Asp	Tyr	Ser	Thr	His	Trp	Asn	Asp	Leu	
143	95					100					105				110		
145	ctt	ttc	tgc	ttg	cgc	tgc	acc	agg	tgt	gat	tca	ggt	gaa	gtg	gag	cta	505
146	Leu	Phe	Cys	Leu	Arg	Cys	Thr	Arg	Cys	Asp	Ser	Gly	Glu	Val	Glu	Leu	
147				115					120					125			
149	agt	ccc	tgc	acc	acg	acc	aga	aac	aca	gtg	tgt	cag	tgc	gaa	gaa	ggc	553
150	Ser	Pro	Cys	Thr	Thr	Thr	Arg	Asn	Thr	Val	Cys	Gln	Cys	Glu	Glu	Gly	
151				130					135					140			
153	acc	ttc	cgg	gaa	gaa	gat	tct	cct	gag	atg	tgc	cgg	aag	tgc	cgc	aca	601
154	Thr	Phe	Arg	Glu	Glu	Asp	Ser	Pro	Glu	Met	Cys	Arg	Lys	Cys	Arg	Thr	
155			145				150					155					
157	ggg	tgt	ccc	aga	ggg	atg	gtc	aag	gtc	ggt	gat	tgt	aca	ccc	tgg	agt	649
158	Gly	Cys	Pro	Arg	Gly	Met	Val	Lys	Val	Gly	Asp	Cys	Thr	Pro	Trp	Ser	
159		160				165					170						
161	gac	atc	gaa	tgt	gtc	cac	aaa	gaa	tca	ggt	ata	aag	cac	agt	ggg	gaa	697
162	Asp	Ile	Glu	Cys	Val	His	Lys	Glu	Ser	Gly	Ile	Lys	His	Ser	Gly	Glu	
163	175					180				185					190		
165	gcc	cca	gct	gtg	gag	gag	acg	gtg	acc	tcc	agc	cca	ggg	act	cct	gcc	745
166	Ala	Pro	Ala	Val	Glu	Glu	Thr	Val	Thr	Ser	Ser	Pro	Gly	Thr	Pro	Ala	
167				195					200					205			
169	tct	ccc	tgt	tct	ctc	tca	ggc	atc	atc	ata	gga	gtc	aca	gtt	gca	gcc	793
170	Ser	Pro	Cys	Ser	Leu	Ser	Gly	Ile	Ile	Ile	Gly	Val	Thr	Val	Ala	Ala	
171				210					215				220				
173	gta	gtc	ttg	att	gtg	gct	gtg	ttt	gtt	tgc	aag	tct	tta	ctg	tgg	aag	841
174	Val	Val	Leu	Ile	Val	Ala	Val	Phe	Val	Cys	Lys	Ser	Leu	Leu	Trp	Lys	
175			225				230					235					
177	aaa	gtc	ctt	cct	tac	ctg	aaa	ggc	atc	tgc	tca	ggt	ggt	ggt	ggg	gac	889
178	Lys	Val	Leu	Pro	Tyr	Leu	Lys	Gly	Ile	Cys	Ser	Gly	Gly	Gly	Gly	Asp	
179		240				245					250						
181	cct	gag	cgt	gtg	gac	aga	agc	tca	caa	cga	cct	ggg	gct	gag	gac	aat	937
182	Pro	Glu	Arg	Val	Asp	Arg	Ser	Ser	Gln	Arg	Pro	Gly	Ala	Glu	Asp	Asn	
183	255				260					265					270		
185	gtc	ctc	aat	gag	atc	gtg	agt	atc	ttg	cag	ccc	acc	cag	gtc	cct	gag	985
186	Val	Leu	Asn	Glu	Ile	Val	Ser	Ile	Leu	Gln	Pro	Thr	Gln	Val	Pro	Glu	
187				275					280					285			
189	cag	gaa	atg	gaa	gtc	cag	gag	cca	gca	gag	cca	aca	ggt	gtc	aac	atg	1033
190	Gln	Glu	Met	Glu	Val	Gln	Glu	Pro	Ala	Glu	Pro	Thr	Gly	Val	Asn	Met	

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/811,088

DATE: 03/30/2001

TIME: 15:03:24

Input Set : A:\07334-324001.TXT

Output Set: N:\CRF3\03302001\I811088.raw

191	290	295	300	
193	ttg tcc ccc ggg gag tca gag cat ctg ctg gaa ccg gca gaa gct gaa	1081		
194	Leu Ser Pro Gly Glu Ser Glu His Leu Leu Glu Pro Ala Glu Ala Glu			
195	305 310 315			
197	agg tct cag agg agg agg ctg ctg gtt cca gca att gaa ggt gat ccc	1129		
198	Arg Ser Gln Arg Arg Arg Leu Leu Val Pro Ala Ile Glu Gly Asp Pro			
199	320 325 330			
201	act gag act ctg aga cag tgc ttc gat gac ttt gca gac ttg gtg ccc	1177		
202	Thr Glu Thr Leu Arg Gln Cys Phe Asp Asp Phe Ala Asp Leu Val Pro			
203	335 340 345 350			
205	ttt gac tcc tgg gag ccg ctc atg agg aag ttg ggc ctc atg gac aat	1225		
206	Phe Asp Ser Trp Glu Pro Leu Met Arg Lys Leu Gly Leu Met Asp Asn			
207	355 360 365			
209	gag ata aag gtg gct aaa gct gag gca gcg ggc cac agg gac acc ttg	1273		
210	Glu Ile Lys Val Ala Lys Ala Glu Ala Ala Gly His Arg Asp Thr Leu			
211	370 375 380			
213	tac acg atg ctg ata aag tgg gtc aac aaa acc ggg cga gat gcc tct	1321		
214	Tyr Thr Met Leu Ile Lys Trp Val Asn Lys Thr Gly Arg Asp Ala Ser			
215	385 390 395			
217	gtc cac acc ctg ctg gat gcc ttg gag acg ctg gga gag aga ctt gcc	1369		
218	Val His Thr Leu Leu Asp Ala Leu Glu Thr Leu Gly Glu Arg Leu Ala			
219	400 405 410			
221	aag cag aag att gag gac cac ttg ttg agc tct gga aag ttc atg tat	1417		
222	Lys Gln Lys Ile Glu Asp His Leu Leu Ser Ser Gly Lys Phe Met Tyr			
223	415 420 425 430			
225	cta gaa ggt aat gca gac tct gcc atg tcc taagtgtgat tctcttcagg	1467		
226	Leu Glu Gly Asn Ala Asp Ser Ala Met Ser			
227	435 440			
229	aagtgagacc ttccctgggtt tacctttttt ctggaaaaag cccaactgga ctccagtcag	1527		
230	taggaaagtg ccacaattgt cacatgaccg gtactggaag aaactctccc atccaacatc	1587		
231	accagtgga tggaacatcc tgtaactttt cactgcactt ggcattatit ttataagctg	1647		
232	aattgtgataa taaggacact atggaaatgt ctggatcatt ccgtttgtgc gtactttgag	1707		
233	atttggtttg ggaatgtcatt gttttcacag cactttttta tcctaattgta aatgctttat	1767		
234	ttattttattt gggctacatt gtaagatcca tctacacagt cgttgtccga cttcacttga	1827		
W--> 235	tactatatga tatgaacctt ttttggggtgg ggggtgcngg gcaattccac tctgtctccc	1887		
236	aggctggagt gcaatggtgc aatcttggct cactatagcc ttgacctctg aggctcaagc	1947		
237	gattctctca cctcagccat ccaaatagct gggaccacag gtgtgcacca ccacgcccgg	2007		
238	ctaatttttt gtattttgtc taaatataag ggctctctat gttgctcagg gtggtctcga	2067		
239	attcctggac tcaagcagtc tgcccacytc agactcccaa agcgggtggaa ttagargcgt	2127		
240	gagcccccat gcttggcctt acctttctac yttttataat tctgtatgtt attattttat	2187		
241	gaacatgaag aaacttttagt aaatgtactt gttacatag ttatgtgaat agattagata	2247		
W--> 242	aacataaaaag gaggagacat acaatggggg aagaagaaga agtccccctgt aagaagttna	2307		
W--> 243	cgntctggtt tccagccttc cctcagatgt actttggctt caatgattgg caacttctac	2367		
244	aggggcccagt cttttgaact ggacaacctt acaagtatat gagtattatt tataggtagt	2427		
245	tgtttacata tgagtcggga ccaaagagaa ctggatccac gtgaagtcct gtgtgtggct	2487		
246	ggtccctacc tgggcagtc cttttgcacc catagccccc atctatggac aggctgggac	2547		
247	agaggcagat gggtagatc acacataaca ataggggtcta tgtcatatcc caagtgaact	2607		
248	tgagccctgt ttgggctcag gagatagaag acaaaatctg tctcccacgt ctgccatggc	2667		
249	atcaaggggg aagagtagat ggtgcttgag aatggtgtga aatggttgcc atctcaggag	2727		

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/811,088

DATE: 03/30/2001

TIME: 15:03:24

Input Set : A:\07334-324001.TXT

Output Set: N:\CRF3\03302001\I811088.raw

```

250 tagatggccc ggctcacttc tgggttatctg tcaccctgag cccatgagct gccttttagg 2787
251 gtacagattg cctacttgag gaccttggcc gctctgtaag catctgactc atctcagaaa 2847
252 tgtcaattct taaacactgt ggcaacagga cctagaatgg ctgacgcatt aagggtttct 2907
253 tcttggtgcc tgttctatta ttgttttaag acctcagtaa ccatttcagc ctctttccag 2967
254 caaaccttc tccatagtat ttcagtcatt gaaggatcat ttatgcagggt agtcattcca 3027
255 ggagtttttg gtcttttctg tctcaaggca ttgtgtgttt tgttccggga ctgggttggg 3087
256 tgggacaaag ttagaattgc ctgaagatca cacattcaga ctggtgtgtc tgtggagttt 3147
257 taggagtgagg gggtgacctt tctgggtcttt gcacttccat cctctcccac ttccatctgg 3207
258 catcccacgc gttgtccctt gcaactcttg aaggcacagg gtgctgctgc ctctggtct 3267
259 ttgcctttgc tgggccttct gtgcaggacg ctcagcctca gggctcagaa ggtgccagtc 3327
260 cggctccagg tcccttgctc cttccacaga ggcttcccta gaagatgcat ctagagtgtc 3387
261 agccttatca gtgtttaaga tttttctttt atttttaatt tttttgagac agaattctac 3447
262 tctctcgccc aggtctggag gcaacggtag gatcttggtc cagtgcacc tccgcctcct 3507
263 gggttcaagc gattctcgtg cctcagcctc cggagtagct gggattgcag gcaccgcga 3567
264 ccacgcctgg ttaatttttg tatttttagt agagacgggg ttccaccatg ttggtcaggc 3627
265 tggctctcga ctcctgacct caggtgatcc accttggcct ccgaaagtgc tgggattaca 3687
266 ggcgtgagcc accagccagg ccaagctatt cttttaaagt aagcttctct acgacatgaa 3747
267 ataattgggg gttttgttgt ttagttacat taggctttgc tatatcccca ggccaaatag 3807
268 catgtgacac aggacagcca tagtatagtg tgtcactcgt ggttggtgtc ctttcatgct 3867
269 tctgccctgt caaaggtccc tatttgaaat gtgttataat acaaacaagg aagcacattg 3927
270 tgtacaaaat acttatgtat ttatgaatcc atgaccaa ataaatatgaa accttatata 3987
271 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gggsgggcgg 4047
272 ccgc 4051
274 <210> SEQ ID NO: 4
275 <211> LENGTH: 440
276 <212> TYPE: PRT
277 <213> ORGANISM: Homo sapiens
279 <400> SEQUENCE: 4
280 Met Glu Gln Arg Gly Gln Asn Ala Pro Ala Ala Ser Gly Ala Arg Lys
281 1 5 10 15
282 Arg His Gly Pro Gly Pro Arg Glu Ala Arg Gly Ala Arg Pro Gly Leu
283 20 25 30
284 Arg Val Pro Lys Thr Leu Val Leu Val Val Ala Ala Val Leu Leu Leu
285 35 40 45
286 Val Ser Ala Glu Ser Ala Leu Ile Thr Gln Gln Asp Leu Ala Pro Gln
287 50 55 60
288 Gln Arg Ala Ala Pro Gln Gln Lys Arg Ser Ser Pro Ser Glu Gly Leu
289 65 70 75 80
290 Cys Pro Pro Gly His His Ile Ser Glu Asp Gly Arg Asp Cys Ile Ser
291 85 90 95
292 Cys Lys Tyr Gly Gln Asp Tyr Ser Thr His Trp Asn Asp Leu Leu Phe
293 100 105 110
294 Cys Leu Arg Cys Thr Arg Cys Asp Ser Gly Glu Val Glu Leu Ser Pro
295 115 120 125
296 Cys Thr Thr Thr Arg Asn Thr Val Cys Gln Cys Glu Glu Gly Thr Phe
297 130 135 140
298 Arg Glu Glu Asp Ser Pro Glu Met Cys Arg Lys Cys Arg Thr Gly Cys
299 145 150 155 160
300 Pro Arg Gly Met Val Lys Val Gly Asp Cys Thr Pro Trp Ser Asp Ile

```

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

**VERIFICATION SUMMARY**

DATE: 03/30/2001

PATENT APPLICATION: US/09/811,088

TIME: 15:03:25

Input Set : A:\07334-324001.TXT

Output Set: N:\CRF3\03302001\I811088.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:235 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:242 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:243 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3

L:463 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5

L:471 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5